9

(51) International Patent Classification 6 : C12N 15/90, 15/55, 9/22, C07K 14/435, A01K 67/027, A61K 48/00	A1	(11) International Publication Number: **WO 99/07871** (43) International Publication Date: 18 February 1999 (18.02.99)

(54) Title: VECTORS AND METHODS FOR PROVIDING CELLS WITH ADDITIONAL NUCLEIC ACID MATERIAL INTEGRATED IN THE GENOME OF SAID CELLS

(57) Abstract

The present invention provides novel elements for improving genetic engineering techniques for producing recombinant nucleic acid molecules and/or recombinant cells. The novel elements are capable of integrating desired necleic acid material into other nucleic acid materials, specifically into the genome of a host cell. The novel elements are derived from or based on transposons, in particular from the Tc/Mariner superfamily. In particular the essential elements of Tc1 enabling excision and pasting of the desired nucleic acid material are provided, together with the relevant transposase activity in cis or in trans.

Title: Vectors and methods for providing cells with additional nucleic acid material integrated in the genome of said cells.

The present invention relates to the field of genetic engineering, in particular to the field of methods for making recombinant vectors containing nucleic acids of interest, methods for making recombinant cells, preferably
5 expressing recombinant products, and/or production of nucleic acids of interest.

In a further embodiment the invention relates to transgenesis of animals, plants or other organisms of medical, scientific or economic interest. In yet another
10 embodiment the invention relates to tools for mutagenesis as well as to tools and means for localization and/or identification of nucleic acid sequences of interest (such as genes) in the genome of a host.

Many ways of providing cells with additional nucleic
15 acids are now known in the field. Many proteins have been expressed in many kinds of cells, ranging from prokaryotic bacteria and bacilli, via yeast and fungi to plant cells, insect cells an mammalian cells.

Often expression of proteins as discussed above has
20 met with success, there are however, certain areas in the field of recombinant technology where, because of safety issues or technical problems, such as stability of the additional nucleic acid material introduced, success has not been easy to achieve. Also, in the application of
25 genetic engineering where the product is not a protein, but for instance a nucleic acid of interest (such as DNA, RNA or an antisense construct), the same or similar problems have been encountered. It is in these areas particularly that the present invention finds its use.
30 One of the problems encountered in recombinant technology is that it is often desirable to have the

2

additional nucleic acid material, which is introduced into
a host cell, integrate in the genome of said host cell.
Once the additional nucleic acid material is integrated in
the genome its stability is much less an issue. Moreover,
5   the integrated material is replicated together with the
genome and thus will be present in the offspring of the
recombinant cell as well. Vectors which are capable of
providing for integration of additional nucleic acid
material into a host cell are known, but they do suffer
10   from some drawbacks which will be discussed below.
        The present invention now provides a way in which a
wide variety of vectors can be altered to integrate the
desired additional nucleic acid in the genome of the host
cell.
15   This means that vectors which did not (efficiently)
integrate the desired additional nucleic acid material in
the genome of the host cell, can now be provided with said
ability.
        One of the drawbacks of the integrating vectors of the
20   prior art is that they are often not capable of transducing
a host cell efficiently. Techniques such as electroporation
and the like are then necessary for achieving transduction.
        In some areas such treatments of cells may be unwanted
or impossible. One such an area is for instance gene
25   therapy. In such cases vectors which can efficiently
transduce host cells on their own, or which can be packaged
into recombinant viral particles and so infect host cells
are often employed.
However, many of these vectors are then incapable of
30   (efficient) integration of the desired additional nucleic
acid material in the genome of the host. The present
invention thus offers the best of both worlds in that it
enables to prepare vectors capable of efficiently
transducing and efficiently integrating desired additional
35   nucleic acid material in a host cell. The inventions solves
this problem in that the nucleic acid to be integrated into
a host cell genome is provided within a functional

transposon. Exogenous DNA has been introduced into the
genome of a host using a transposon, however since
transposons were, until the present invention, thought to
be rather species specific, the applicability of such a
5   system was thought to be very limited. At best, it was
shown that a transposon functional in one fruit fly could
also be used to integrate DNA into the genome of another
species of fruit fly (Loutheris et al., 1995). The present
inventors have found that, at least for a certain class of
10   transposons, it is possible to use these transposons across
wide phylogenetic barriers, which allows for wide
application of these transposons as integration means for
DNA into a wide variety of host genomes.
The invention thus provides a vector for providing a cell
15   of a certain genus with additional nucleic acid material
integrated in its genome, whereby said vector comprises two
transposase binding sites, whereby said transposase binding
sites may be the same or different and are derived from a
transposon found in another genus, each in close proximity
20   to a cut site for said transposase, whereby said additional
nucleic acid material is located between said two
transposase binding sites. It is of course clear that it is
highly preferred to be able to use transposons over even
wider phylogenetic barriers. The present invention thus
25   provides a transposon-based integration system with a wide
applicability.
          Clear advantages of this system are for instance that:
using a transposon will often lead to a more efficient
integration into the host cell genome; using a transposon
30   gives complete control of the integrating sequence (the
termini of the integrating element are known); in the
embodiment where gene (or DNA) localization is desired it
will be an advantage that a transposon can integrate
anywhere in the genome (especially useful in so-called
35   mutagenesis experiments aimed at gene function, or in
"gene-trap" experiments directed at tracing genes of
interesting expression patterns). For expression purposes

it may be desirable to develop transposons with a more
specific integration site in the genome, or to develop
methods of integration which lead to a more specific
integration site.

5   Preferably, said transposase binding sites and said
cut sites are based on the corresponding sites in
transposons from the Tc1/mariner superfamily of
transposons, more preferably they are based on Tc1-like
transposons. A very useful set of transposon elements is

10   the minimum set required by the Tc1 transposon i.e.
comprising the terminal 26 basepairs of Tc1 and in close
proximity the cut site (TA) for Tc1 transposase.

        An important advantage of the transposons according to
the invention is that they are self-sufficient. All that is

15   needed is a functional transposase binding site, a
transposase cut site and transposase activity functional
for those sites. Transposase activity needed should be as
limited as possible. The preferred class of transposons
only needs a single transposase. A reason for the species-

20   specificity of transposons reported until to date may be
that transposons have been used that require host proteins
in order to be able to jump. The host proteins may then be
responsible for the species-specificity of for instance the
transposons of bacterial origin, or the P-element. The

25   class provided with the present invention needs no host-
elements and has modest cis-trans requirements. Therefore
these transposons are far more suitable for integrating
nucleic acids of interest into the genome of a wide variety
of hosts.

30   In a preferred embodiment where, elements based on Tc1
are used such a transposon binding site seems to be at
least the 26 terminal basepairs. Probably not all of these
basepairs are essential for the function of this binding
site. Some (conservative) changes in such a binding site

35   may therefore be allowed. It seems that sequences of around
100 bp are most efficient as transposase recognition sites.

For the invention, the transposase recognition sites
must be functional for the corresponding transposase-
activity, which transposase of course may also be modified,
mutated, shortened or lengthened when compared with the
5   original transposase, as long as it has relevant
transposase-activity. The transposase-activity may be
encoded on the same vector as the other transposon
elements, even as part of the transposon together with the
desired additional nucleic acid material (in cis), but it
10   may very suitably be provided to the cell to be transduced
by another vector according to the invention or by another
vector (in trans).

When the transposase-activity is provided in trans and
is preferably encoded by a sequence under control of an
15   inducable and/or repressible promoter, "jumping" of the
transposon can be switched on and off. Once integration
into the host genome has taken place this may be very
useful.

A vector for the purposes of this invention is any
20   nucleic acid vehicle which is capable of carrying the
desired additional nucleic acid material and preferably
capable of transduction of host cells and/or replication.

The desired additional nucleic acid material may be
encoding a protein and thus comprise a gene and regulatory
25   elements for expression. Proteins to be expressed will
depend on the purpose of the transduction. Many useful
proteins for many applications have been identified as good
candidates for recombinant expression. All these may be
expressed using the present invention.

30   Also, the additional nucleic acid material (DNA) may
be transcribed (once transduced) into RNA molecules
blocking transcription or translation of host cell (or
viral) nucleic acids. Such additional nucleic acid material
may for instance be an antisense RNA molecule.

35   Use of the invention in the field of gene therapy is
particularly advantageous, as will be clear from the above.
Viral vectors often contemplated for gene therapy (such as

6

adenovirus, retrovirus or adenoassociated virus) which are
not capable of efficiently integrating the desired nucleic
acid material into the genome of the infected cell, may now
be provided with such capability. Other vectors not based

5   on a virus, but provided with a means to deliver them to a
target cell population can of course also be advantageously
provided with an integration system according to the
invention. Particularly the use of liposomes or polymers as
a targetting system is contemplated.

10   As stated, the present invention provides an efficient
means by which a variety of vectors may be used to
integrate exogenous nucleic acid material in the genome of
a host cell. Because the exogenous nucleic acid material is
integrated into the host cell's genome the progeny of the

15   recombinant host cell produced will possess the same
recombinant capabilities as their parent. Thus, when
stemcells are produced according to the invention, this
will lead to very efficient gene therapy regimes. The
vectors (or the integration system) of the invention can of

20   course also be used to produce transgenic animals, plants
or other organisms of scientific, economic or medical
interest. Useful applications for transgenesis are well
known in the art.

        Thus methods using the vectors according to the

25   invention to transduce cells or animals, plants, etc. are
also disclosed and part of the invention, as are cells,
animals or plants, etc. obtainable by such methods. Use of
the invention in mutagenesis studies and the like is
another preferred embodiment, as discussed hereinbefore.

30   The invention will now be explained in more detail using
Tc1 as a non-limitative example.

Detailed description

35   Tc1 belongs to the Tc1/mariner superfamily of
transposons found in nematodes, arthropods and chordates

(Henikoff 1992; Raddice et al. 1994; Robertson 1995;
Plasterk 1995). Both vertical and horizontal transfer have
contributed to the spread of these elements throughout the
animal kingdom (Robertson 1993; Radice et al. 1994;
5   Robertson and Lampe 1995). The widespread occurrence of the
Tc1/mariner family of transposons can be taken as an
indication for the absence of species-specific host factors
which limit the transfer between different species.
Therefore, Tc1/mariner elements are attractive candidates
10   for the development of gene delivery vectors.
        Tc1-like elements are close to 1.7 kb in length, have
short inverted terminal repeats flanking a transposase gene
and have the conserved sequence CAGT at their termini,
flanked by TA representing the target site, which is
15   duplicated upon integration (Van Luenen et al. 1994). The
element-encoded proteins share a homologous catalytic
domain with bacterial transposases and retroviral
integrases (Doak et al. 1994). Tc1 from *C. elegans* is a
1612 bp long transposon which has 54 bp inverted repeats
20   flanking a gene encoding a 343 amino acid transposase
(Emmons et al. 1983; Rosenzweig et al. 1983; Vos et al.
1993), that binds to the inverted repeats (Vos et al. 1993;
Vos and Plasterk 1994). The conserved hexanucleotide
sequence, TACAGT (SEQ ID NO: -), at the extreme termini of
25   the element is not part of the transposase binding site,
but is thought to play a role in catalysis of the
transposition reaction (Vos and Plasterk 1994).
        Here, we describe <u>in vitro</u> excision and transposition
of Tc1 using an extract prapared from transgenic nematodes.
30   The minimal cis-requirements for transposition are defined
and the target site choice in vitro is compared with that
<u>in vivo</u>. Furthermore, we demonstrate that recombinant
transposase purified from *E. coli* is capable of supporting
transposition, showing that no other factors are essential
35   for Tc1 transpostion <u>in vitro</u>.

8

Moreover, we show that Tc1-transposase supports
jumping of Tc1-derived transposons carrying a neomycin
resistance gene into the genome of human cells. This
demonstrates that it is possible to construct transposon-

5   derived vectors that are capable of providing for
integration of additional nucleic acid material into host
cells across large phylogenetic barriers, which makes such
vectors suitable as integrating gene delivery vehicles for
a wide variety of host genomes, including that of humans.

10

Results

Transposition of Tc1 in vitro

We generated a transgenic worm with the Tc1
transposase gene under the control of a heat shock

15   promoter. This allowed the preparation of a nuclear extract
with elevated levels of transposase, which proved to be
highly important to detect activity. The extract was
incubated with a plasmid containing a Tc1 element. Excision
was studied in a physical assay. Southern blot analysis of

20   reaction products shows the appearance of excised Tc1
elements (fig. 1, lane 1).

Furthermore, cleavage at either the left or the right
end of Tc1 is detected when the products are digested with
ScaI within the plasmid backbone prior to electrophoresis

25   (lane 2). Cleavage may require a divalent cation (Mg^{2+} or
Mn^{2+}) and is stimulated by the presence of ethylene glycol
or 5% DMSO (data not shown). The efficiency of cleavage at
a single end of the transposon is not decreased if the
substrate is linear (compare lanes 2 and 5). Also, deletion

30   of either end of the transposon does not abolish cleavage
at the remaining end (lanes 6 and 11), which suggests that
cleavage does not require interaction between the two ends.
In contrast to single end cleavage, excision of the
complete element is reduced about 2-fold when the substrate

35   is linear, suggesting that coordinated cutting at both ends
is stimulated by supercoiling of the substrate. The

majority of complete excision products observed with a
linear substrate can be explained by non-coordinated
cleavages at both ends.

5   To determine the positions of the double strand
cleavages at the nucleotide level, a PCR based primer
extension was performed using end-labeled oligonucleotides
specific for each strand (Fig. 2). The 5' cut is 2 bp
within the transposon, whereas the 3' cut maps to the end
of the transposon, as based on the largest observed PCR
10   product. This confirms the model based on _in vivo_ studies
of the related _C. elegans_ transposon Tc3, for which it was
shown that excision results in a 2 bp staggered 3' overhang
(Van Luenen et al. 1994). The complete excision of Tc1
shows that transposition occurs via a cut-and-paste
15   process, a result consistent with genetic data on double
strand break repair of the donor DNA molecule upon Tc1
excision (Plasterk 1991).

We devised a sensitive assay to detect integration
events. We selected for jumping of a transposon-borne
20   antibiotic resistance gene from a supercoiled donor plasmid
to a target plasmid in a genetic assay (Fig. 3).
Electroporation of reaction products into the appropriate
E. coli strain resulted in the detection of many
transposition events (Table 1). Extracts prepared from non-
25   transgenic N2 worms or from the so-called high-hopper
strain, TR679 (Collins et al. 1987), which has a high
frequencey of germline transposition, do not generate a
detectable level of transposition products in this assay.
Linearization of the donor plasmid resulted in an
30   approximately 20-fold reduced efficiency of transposition.
Transposition requires two inverted repeat sequences,
because no integrations were obtained upon deletion of one
transposon end. Furthermore, the addition of ATP, GTP or
dNTPs does not increase the level of transposition (data
35   not shown), which indicates that the process is neutral in
energy-consumption and independent of a cofactor. About 90

independent <u>in vitro</u> Tc1 integrations were analyzed by
sequencing and found at TA dinucleotides, which had been
duplicated in the process. Two odd integration events were
detected, where Tc1 had integrated in the sequence TTG (SEQ
5   ID NO: -) or CCT (SEQ ID NO:-). In both cases, we found a 3
bp target site duplication.

Target site choice
        Previously, several hundreds of <u>in vivo</u> Tc1 and Tc3
10   integrations in a 1 kb region of the *gpa-2* gene have been
analyzed (Van Luenen and Plasterk 1994). This showed the
selective use of a limited set of TA dinucleotides as
targets of integration, with a striking difference in
preference between Tc1 and Tc3. To investigate whether the
15   chromatin structure played a role in the choice of
integration sites, we determined the pattern of
integrations into naked DNA <u>in vitro</u>, using the same target
region previously assayed <u>in vivo</u>. Therefore, we included
the *gpa-2* region in our target plasmid. It is apparent that
20   the same overall pattern of integration is seen (Fig. 4).
Hot sites <u>in vivo</u> appear to be hot in vitro and cold sites
in vitro are also cold <u>in vivo</u>. This indicates that, at
least in this region of the genome, the genome, the
chromatin structure or the transcriptional status of the
25   DNA <u>in vivo</u> is not the major determinant of target choice.

Transposition by recombinant transposase
        The nematode is not a convenient source of protein for
an extensive purification of transposase. Therefore, we
30   expressed the protein in a heterologous system. Both
expression using Baculovirus and Sf9 cells (data not shown)
or expression in *E. coli* yielded transposase capable of
supporting Tc1 transposition. Recombinant transposase was
purified from inclusion bodies to near homogeneity (Fig.
35   5). Table 1 shows the frequency of transposition when
comparable amounts of transposase were used for both the

worm extract and the purified protein. Sequence analysis of
9 independent integrations in case of the recombinant
protein showed that transposition into TA target sequences
that were duplicated, from which it can be concluded that
5   bona fide transposition had occurred. Therefore, we
conclude that Tc1 transposase is the only protein required
for Tc1 transposition. The difference in efficiency between
nematode derived and bacterial transposase needs further
studies. It could reflect a folding problem of the
10 bacterial transposase, which was denatured and refolded
during the purification procedure, or the stimulatory role
of host factors present in the nematode extract.

Minimal cis-requirements

15   We investigated the possibility that the terminal 26
bp of Tc1 which constitute a full transposase binding site,
flanked by the TA target site, are sufficient to form an
artificial transposon. An element consisting of only these
Tc1-specific sequences is still able to transpose in vitro,
20 albeit at a lower frequency (Table 1). We sequenced several
integrations and found them to be correct.
       Furthermore, we investigated the importance of the
conserved hexanucleotide sequence TACAGT (SEQ ID NO.: -).
Mutations were introduced at one of the ends of a mini-Tc1
25 which contains only the terminal 26 bp as well as the
flanking TA dinucleotide. Whereas excision of the element
with 2 wild-type ends is easily detected in a physical
assay, mutation of the transposase binding site, the
flanking TA sequence of the termini, resulted in the
30 inability of the element to excise (Fig. 6). Double strand
cleavage at the wild-type end was not affected by mutation
at the other transposon end. Analysis of cleavage by PCR
based primer extension revealed that, for the CA to TG
mutation only, single stranded breaks at the 5' end of the
35 transposon had occurred (data not shown).

We have developed a cell-free *Tc1* transposition
system. Excision occcurs by double strand breaks at the
transposon ends resulting in 2 bp staggered 3' overhangs. A
cut-and-paste mechanism of transposition appears to apply

5   for *Tc1* (Fig. 7). This mechanism was already proposed on
the basis of genetic data (Plasterk 1991) as well as the
analysis of <u>in vivo</u> transposition products (Van Luenen et
al. 1994). Nonreplicative transposition is shared with the
bacterial transposons Tn7 (Bainton et al. 1991, 1993) and

10 Tn*10* (Bender and Kleckner 1986) as well as the *Drosophila* P
element (Kaufman and Rio 1992). In contrast, the Mu and Tn*3*
transposable elements transpose via a replicative mechanism
(Grindley and Sherratt 1978; Shapiro 1979; Mizuuchi 1992).
Tc1 transposition appears to be independent of addition of

15 a nucleotide cofactor, whereas P elements use GTP (Kaufman
and Rio 1992) and Tn7 uses ATP as cofactor (Bainton et al.
1993).

A striking feature of the *Tc1/mariner* family is the
use of a TA dinucleotide as target site. An extensive study

20 of target site choice <u>in vivo</u> had revealed the usage of
only a subset of the available TA dinucleotides and a
marked difference in target choice between the two related
transposons *Tc1* and *Tc3* in *C elegans* (Van Luenen and
Plasterk 1994). We find the same overall integration

25 pattern <u>in vitro</u> as had been observed <u>in vivo</u>. This
suggests that the chromosomal context of the DNA does not
affect target choice, at least in the region of the genome
analyzed. Therefore, we favor the idea that the
transposition complex primarily selects its target site on

30 the basis of the primary DNA sequence flanking the TA,
although a strong consensus sequence could not be
identified (Van Luenen and Plasterk 1994). A clear
influence of the chromatin structure has been demonstrated
for retroviral integrations (Pryciak and Varmus 1992;

35 Müller and Varmus 1994). These studies showed a preference
for regions within nucleosomal DNA, probably due to the

bending of the DNA. We cannot exclude that DNA binding
proteins can affect regional preferences for Tc1
integration. Because nothing is known about the chromosomal
organization of the gpa-2 gene, it will be of interest to
5   compare integration sites using reconsituted nucleosomal
DNA in vitro.

Transposition in vitro requres the extreme termini of
the transposon containing the transposase binding site and
the conserved hexanucleotide sequence, which is important
10   for excision. We observe a decrease in transposition
efficiency between transposition of a full-length
transposon and the Tc1 element with only 26 bp terminal
inverted repeats, which suggests that additional sequences
can contribute to transposition efficiency. We have no
15   indications for additional transposase binding sites, but
perhaps small basic proteins like high mobility group
proteins (Grosschedl et al. 1994) may bind and stimulate
transposition. Alternatively, unique A-T-rich sequences
found at the transposon ends may add a helping bend to the
20   DNA. The conserved hexanucleotide sequence at the extreme
termini of the transposon are shown to be important at
least for the cleavage step. The 5' end single strand
cleavage seen for one of the mutations (CA to TG) is
perhaps an indication for a specific order of single strand
25   cleavages, i.e. first the non-transferred strand, which
would be the opposite of what has been reported for Tc10
(Bolland and Kleckner 1995).

Transposase purified from E. coli to near homogeneity
is able to execute jumping of Tc1, which indicates that
30   transposase is the only protein required for excision and
integration of Tc1. The higher efficiency obtained with the
nematode extract suggests that host factors may enhance the
frequency of the reaction. It has for instance been shown
that the mammalian proteins HMG1 and HMG2 can stimulate
35   prokaryotic recombinations (Paull et al. 1993). The
independence of species-specific factors might be the

explanation why members of the Tc1/mariner family are
dispersed over so many different phyla, possibly by means
of horizontal transfer (Robertson and Lampe 1995). This is
in contrast to P elements which are restricted to

5   Drosophila species. Transposition of P elements in other
species has not been observed (Rio et al. 1988). A possible
candidate for a species-specific host factor in
P transposition is the inverted repeat binding protein,
IRBP (Beall et al. 1994). The simple cis- and trans-

10   requirements for Tc1 transposition in vitro shows that this
transposable element will be a good vector for gene
delivery in a wide variety of animals. In order to
determine the feasibility of the use of Tc1 transposons in
human cells, we first tested whether the Tc1-transposase is

15   toxic for human cells. The Tc1-transposase expression
vector pRc/CMV.Tc1A or the empty vector pRc/CMV
(Invitrogen) was transfected into 911 cells, a human embryo
retina cell line transformed by early region I of
adenovirus type 5. Approximately 48 hours post-

20   transfection, the cells were put on G418 selection medium.
Stable cell lines derived from single colonies were
established and screened for Tc1-transposase protein on a
western blot (figure 8). All pRc/CMV.Tc1A transfected cell
lines expressed Tc1-transposase, demonstrating that

25   expression of Tc1-transposase is not toxic for human cells.
Next, we examined whether Tc1-transposase supports
transposition of a Tc1-derived transposon in human cells.
To that end, the plasmid pRP466.SV-neo1.PGK-tk2 was
constructed. This plasmid contains a neomycin resistance

30   gene expression cassette under control of the SV40 promoter
(SV-neo) flanked by the inverted repeats of Tc1and an
expression cassette for the Herpes Simplex Virus thymidin
kinase under control of the phospho-glyceraldehyde kinase
promoter (PGK-tk) outside of the Tc1 inverted repeats.

35   Excision of the SV-neo transposon from pRP466.SV-neo1.PGK-
tk2 was tested in an in vitro excision assay using

transposase purified from E.coli.. Southern blot analysis
of the reaction products shows the appearance of excised
2.9 kb Tc1.SV-neo elements (figure 9). Thus, pRP466.SV-
neo1.PGK-tk2 can serve as a Tc1.SV-neo transposon donor
5   plasmid.

     To determine whether transposition of Tc1 transposon-
derived vectors occurs in human cells, the cell line 911
was transfected with the plasmid pRP466.SV-neo1.PGK-tk2,
together with the Tc1-transposase expression vector
10   pcDNA1/Tc1A or with the empty vector pcDNA1/amp
(Invitrogen). Fourty eight hours post-transfection, the
cells were put on G418 selection. After 18 days, G418
resistant colonies were counted (table 2). Co-transfection
of pcDNA1/Tc1A yielded approximately 40% more G418
15   resistant colonies, suggesting that , in addition to random
integration, G418 resistant colonies were formed due to
jumping of the Tc1-neo transposon into the cellular genome.

     The presence of the PGK-tk expression cassette on the
plasmid pRP466.SV-neo1.PGK-tk2 outside of the Tc1 inverted
20   repeats allowed us to discriminate between plasmid
integration and transposition of the Tc1.SV-neo transposon.
Thymidin kinase phosphorylates the anti-viral prodrug
Ganciclovir (GCV), after which it acts as a chain
terminator of DNA synthesis, thereby killing the tk
25   expressing cells. When plasmid integration occurs, in most
cases both the SV-neo and PGK-tk expression cassettes are
integrated into the 911 genome, rendering the cell lines
G418 resistant and GCV sensitive. When transposition of the
SV-neo transposon takes place, however, the cell lines will
30   be G418 resistant and GCV insensitive, because the PGK-tk
expression cassette is not integrated into the host cell
genome. In total, 73 independent G418 resistant cell lines
were established from the above mentioned transfection, and
screened for to GCV. Table 3 shows that co-transfection of
35   pRP466.SV-neo1.PGK-tk2 with pcDNA1/Tc1A resulted in the

16

formation of more tk-negative cell lines (64-77%) than co-
transfection with pcDNA1/amp (32-44%).

Random integration of circular plasmid DNA first
requires linearization of the DNA due to a double strand
5   DNA break before integration into the host genome can
occur. The PGK-tk expression cassette covers approximately
1/3 of the total of the plasmid DNA (excluding the neomycin
SV-neo expression cassette for which the cell lines were
selected). If this process is assumed random and if a
10   single copy of the plasmid integrates, one would expect
around 33% of the DNA breaks to occur in the PGK-tk
expression cassette, thereby abolishing tk expression.
Therefore, the percentage tk negative cell lines that were
generated by co-transfection of pRP466.SV-neo1.PGK-tk2 and
15   pcDNA1/amp was within expectation. The higher incidence of
GCV resistant clones after co-transfection of pRP466.SV-
neo1.PGK-TK2 with pcDNA1/Tc1A again strongly suggests that
the Tc1.SV-neo transposon jumped from the pRP466.SV-
neo1.PGK-TK2 into the human genome.
20   The most direct way to discriminate between plasmid
integration- and transposition events is sequencing the DNA
that flanks the transposon that is integrated into the host
cell genome. After plasmid integration, the surrounding DNA
sequences will be identical to the sequences flanking the
25   transposon in the plasmid pRP466.SV-neo1.PGK-tk2. After
transposition, the surrounding sequences will be of human
genomic origin, and thus be different from those in the
plasmid pRP466.SV-neo1.PGK-tk2. Genomic DNA from
independent tk-minus/G418-resistant colonies was isolated
30   and digested with Sau3A. To the genomic fragments, an
oligo-cassette ("vectorette") was ligated. The vectorette
consists of two oligonucleotides which have complementary
sequences at their ends but are not complementary in the
middle. Fragments that contain transposon and flanking
35   sequences are amplified in a PCR using an oligo
complementary to the transposon end and pointing to the
flanking sequence in combination with an oligo that is

17

complementary to middle of the vectorette sequence. PCR
products were analysed on an agarose gel and the amplified
DNA fragments were sequenced. Thus far, in two out of five
cases we found that the transposon was present, flanked by
5   a TA dinucleotide and followed by a new sequence. These
integrations represent genuine transposition events showing
that Tcl can jump in human cells when Tcl-transposase is
provided.

10   **Materials and methods**
Plasmid constructions

     pRP466 contains a Tcl element with 0.4 kb flanking
sequences derived from plM40 (Mori 1988) cloned as a *BamHI-
XbaI* fragment into pUC19, pRP467 and pRP468 are derivatives
15   of pRP466 in which either a *ClaI-Asp718* or *PstI-ApaI*
fragment is deleted. pRP472 is a pACB104 (Boyd and Sherratt
1995) derivative which contains Tcl with the *AvaI-HindIII*
fragment of pBR322 inserted between the *ClaI* and *ApaI*
sites. Cloning of the *XbaI-BamHI* fragment of pRP466 with
20   the *HindII* Kan^R-cassette of pUC4K (Pharmacia) between the
XhoI-sites into pACB104 resulted in pRP490. pRP491 is
comparable to pRP490; all the internal Tcl sequences have
been replaced except the terminal 26 bp.
     The plasmid pRP466.SV-neol.PGK-tk2 was constructed as
25   follows: pRP466 was digested with *XhoI* and blunted with
Klenow and dNTPs. The plasmid pRc/CMV (Invitrogen) was
digested with *BamHI* and *EcoRI* and blunted with Klenow and
dNTPs. Subsequently, the *BamHI/EcoRI* SV-neo fragment was
cloned into *XhoI* digested pRP466, giving rise to pRP466.SV-
30   neol. The plasmid pRP466.SV-neol was digested with *EcoRI*
and blunted with Klenow and dNTPs. The plasmid pPGK-tk
(PCT/NL/00195) was digested with *PvuII*. The *PvuII* PGK-tk
fragment was cloned into *EcoRI* digested pRP466.SV-neol,
giving rise to pRP466.SV-neol.PGK-tk2.

18

The plasmid pRc/CMV.Tc1A contains the Tc1-transposase
cDNA, which was amplified from pRP470 (Vos *et al.*, 1993)
with the primers
5'CCCCAAGCTTGCCACCATGGTAAAATCTGTTGGGTGTAAAAATC (SEQ ID

5   NO.:-) and   5'GCTCTAGATGCTTAATACTTTGTCGCGTATCC
(SEQ ID NO:-) using eLONGase according to standard protocol
of the supplier (Gibco BRL). PCR was performed on a
Biometra Trio Thermoblock, Amplification program: 94°C for
1 minute, 1 cycle; 94 °C for 30 seconds + -52 °C for 30

10   seconds + 68 °C for 1 minute , 35 cycles; 68 °C for 1
minute. The PCR fragment was digested with HindIII/*Xba*I and
cloned into HindIII/*Xba*I digested pRc/CMV, giving rise to
pRc/CMV.Tc1A. The integrety of the Tc1A cDNA was confirmed
by sequencing. The pRc/CMV plasmid contains a neomycin

15   resistance gene expression cassette. The plasmid
pcDNA1/Tc1A contains the HindIII/*Xba*I Tc1A fragment of
pRc/CMV.Tc1A cloned into HindIII/*Xba*I digested pcDNA1/amp
(Invitrogen). The pcDNA1/amp plasmid does not contain a
neomycin resistance gene expression cassette. All

20   restriction enzymes, primers, Klenow and dNTPs were
purchased from Gibco BRL.

Southern blotting

For figure 9: The product of the *in vitro* excision

25   reaction was separated on a 0.8% agarose 0.5 x TBE gel by
electrophoresis for 5 hours at 70 Volts. Subsequently the
gel was rinsed in Southern blot buffer (0.4 M NaOH, 0.6 M
NaCl) for 30 minutes and the DNA was blotted overnight onto
Hybond-N⁺ (Amersham). After blotting, the gel was rinsed

30   in 5xSSPE and pre-hybridised for 3 hours in 50% formamide,
5xDenhardts, 5xSSPE, 5% dextrane sulphate, 1% SDS and 200
µg/ml haring sperm DNA and overnight hybridised with a
randomly primed ³²P-labelled (RTS radprime DNA labeling
system, Gibco BRL) neo probe (*Nco*I fragment of pPGK-tk) at

35   42°C. Subsequently the blot was extensively washed in

5xSSPE/0.1% SDS at 65°C. The blot was exposed to HyperFilm
(Amersham) at room temperature.

Western blotting

5   Cells were washed twice with PBS (NPBI) and lysed and
scraped in RIPA (1% NP-40, 0.5% sodium deoxycholate and
0.1% SDS in PBS, supplemented with 1mM
phenylmethylsulfonylfluoride and 0.1 mg/ml trypsin
inhibitor). After 15 minutes incubation on ice, the lysates
10   were cleared by centrifugation.Protein concentrations were
determined by the Bio-Rad protein assay, according to
standard procedures of the supplier (BioRad). Whole-cell
extracts were fractionated by SDS-PAGE on 10% gels.
Proteins were transferred onto Immobilon-P membranes
15   (Millipore) and incubated with an αTc1 transposase antibody
(Schukkink *et al.*, 1990). The secondary antibody is a
horseradish-peroxidase conjugated goat anti rabbit antiboby
(BioRad). The antibody complexes were visualized with the
ECL detection system according to the manufacturer's
20   protocol (Amersham).

Cell culture and Transfection

        The cell line 911 (Fallaux *et al.*, 1996) was cultured
in Dulbecco's modified Eagles medium (Gibco BRL)
25   supplemented with 10% FBS (Gibco BRL) at 37°C and 5% CO_2.
Selection of the transfected cells with G418 (Gibco BRL)
was performed at a concentration of 500 µg/ml. Selection
with Ganciclovir (Roche) was performed at a concentration
of 10 µg/ml. DNA was transfected into 911 cells using the
30   Calcium Phosphate Transfection System, according to the
manufacturer's protocol (Gibco BRL). Briefly, 1 day prior
to transfection, 911 cells were plated at 30% density on 5
cm tissue culture dishes (Greiner). The next day, cells
were incubated with calcium phosphate precipitated DNA for
35   15 hours. Subsequently, the cells were washed twice with
PBS and frsh medium was added. Approximately 48 hours post-

transfection cells were put on G418 selection medium. G418
resistant colonies appeared around day 6. Single colonies
were picked and cell lines were established.

5   *Sequencing of the transposon flanks*

         100 ng genomic DNA was digested in 20 µl using 4 units
     Sau3A (Boehringer Mannheim) according to instructions of
     the supplier. After 2 hours at 37°C the enzyme was
     inactivated by incubating 15 minutes at 65°C. Equimolar
10   amounts of the vectorette oligo's (503:
     5'GATCCAAGGAGAGGACGCTGTCTGTCGAAGGTAAGGAACGGACGAGAGAAGGGAGA
     and 504
     5'TCTCCCTTCTCGAATCGTAACCGTTCGTACGAGAATCGCTGTCCTCTCCTTG)
     were mixed in the presence of 10 mM Tris-HCl pH7.5 and 1mM
15   EDTA. The mixture was heated to 80°C and allowed to cool
     down to 40°C. The final concentration of the vectorette-
     oligo-cassette is 10 pmol/µl. 15 pmoles of the vectorette
     were overnight ligated at 16°C to the digested DNA in a 100
     µl volume using 5 units ligase (Boehringer Mannheim)
20   according to the instructions of the supplier. In the first
     PCR, 3 µl of the DNA is used in a volume of 25 µl using 1
     unit Taq polymerase (Gibco BRL) according to the
     instructions of the supplier. The final concentration of
     the oligo's is 0.4 pmoles. The PCR consisted of 30 cycles
25   with 1 minute at 95°C, 1 minute at 58°C and 1 minute at
     72°C. To increase the specificity and the sensitivity a
     second PCR with nested primers was performed using 0.01 µl
     template from the first PCR. The conditions of the PCR were
     identical to the first PCR. The oligo's used are Tc1L2 (5'-
30   TCAAGTCAAATGGATGCTTGAG) (SEQ ID NO:5) or Tc1R2
     (5'- GATTTTGTGAACACTGTGGTGAAG) (SEQ ID NO:6) and 337new
     (5'-GTACGAGAATCGCTGTCCTC) (SEQ ID NO:7). The PCR products
     were analysed on a 1 % agarose gel. Amplified DNA fragments
     were excised from gel and DNA was extracted using the QIAEX
35   II gel extraction kit (Qiagen). The DNA was taken up in 20
     µl water after isolation from gel. The transposon primer

Tc1L2 or Tc1R2 was radiolabeled using ^{32}P-ATP and polynucleotide kinase (Boehringer Mannheim) according to the instructions of the supplier. Using 0.25 pmoles of radiolabeled primer and the isolated DNA a PCR was

5   performed using 0.5 units Taq polymerase (Gibco BRL) and one of the ddNTP mixtures (containing either ddATP (160 μM ddATP and 5 μM of each dNTP in 10 mM Tris-HCl pH 8, 1 mM EDTA), ddTTP (250 μM ddTTP and 5 μM of each dNTP in 10 mM Tris-HCl pH 8, 1 mM EDTA), ddGTP (32 μM ddGTP and 5 μM of

10   each dNTP in 10 mM Tris-HCl pH 8, 1 mM EDTA) or ddCTP (160 μM ddCTP and 5 μM of each dNTP in 10 mM Tris-HCl pH 8, 1 mM EDTA)). The final volume was 20 μl. 20 cycles were performed with 1 minute at 95°C; 1 minute 58°C; 1 minute 72°C. The samples were analysed on a 6 % denaturing

15   polyacrylamide gel in 1 x TBE.

Transgenesis of C. elegans

       A transgenic Bristol line was obtained after microinjection (Mello et al. 1991) of 150 mg/ml pRP469 and

20   5 mg/ml pRP465 (Vos et al. 1993), 50 mg/ml pRF4 (Kramer et al. 1990) in strain CB1392 (nuc-1 9e1392)). A stable line, NL818(pk1s221), was generated by X-ray irradiation (Way et al. 1991).

25   *Extract preparation*

       Stable line NL818 was grown in liquid culture at 18°C and heat shocked for 3 hours at 33°C to induce transposase expression. After 2 hours of further growth at 18°C, nuclear extracts were prepared as described (Vos et al.

30   1993) with differences in the buffers. NIB: 25 mM Tris pH 7.5, 20 mM KCl, 0.5 M sucrose, 0.5 mM EDTA, 5 mM b-mercaptoethanol, 0.1 mM PMSF, NEB: 25 mM Tris pH 7.5, 0.1 mM EDTA, 500 mM NaCl, 15% glycerol, 0.25% Tween-20, 0.1 mM PMSF, 1 mM DTT. Nuclear extract contains 2.5 mg/ml protein;

35   concentration of Tc1A is about 10 mg/ml.

22

Recombinant transposase expression and prurification

     E. coli strain BL21 pLysS was transformed with pRP470
containing the Tc1 transposase gene under the control of a
T7 promoter (Vos et al. 1993), grown in 2xYT medium and
5   induced at an OD of 0.6 at 600 nm with 0.5 mM IPTG for 3
hours at 37°C. Inclusion bodies were purified as described
by Nagai and Thogersen (1978). Inclusion bodies were
dissolved in 8 M urea, 20 mM Na-phosphate pH 6.0 and
loaded on a CM cellulose CL-6B column (Pharmacia). The
10   protein was eluted with a linear gradient from 0 to 500 mM
NaCl. The transposase containing fraction was loaded on a
Sephacryl S400 HR gel filtration column equilibrated in 6 M
guanidiumhydrochloride, 50 MM Tris pH 8.0. Transposase
fractions were dialysed against 8 M urea, 50 mM Tris pH
15   8.0, 1 mM DTT. The protein was loaded on a S Sepharose FF
column and eluted with 500 mM NaCl in the same buffer. All
steps were performed at room temperature. The protein was
renatured by a 100xdilution into ice-cold buffer: 50 mM
Tris pH 8.0, 100 mM NaCl, 5 mM DTT, 5 mM $MgCl_2$. After
20   30 minutes, insoluble protein was removed by centrifugation
for 15 minutes in an Eppendorf centrifuge. Transposase
concentration was 200 mg/ml and estimated to be more than
90% pure.

25   *In vitro transposition reactions*

     Standard reaction conditions: 25 mM Tris pH 8.0, 25 mM
NaCl, 1 mM DTT, 10% ethylene glycol, 5 mM $MgCl_2$ (or 2.5 mM
EDTA), 4 mM spermidine, 0.05 mg/ml BSA. 200 ng of donor
plasmid was preincubated with 2.5 ml worm extract or
30   0.25 ml of purified protein for 5 minutes on ice before
addition of 2.5 mg target DNA in a total volume of 50 ml.
Incubation was for 1 hour at 30°C. Reactions were stopped
by addition of 5.5 ml of 250 mM Tris pH 8.0, 50 mM EDTA, 5%
SDS, 2 mg/ml proteinase K. After 1 hour at 37°C, the DNA
35   was precipitated and resuspended in 50 ml water.

Mapping of in vitro cleavage sites

Linear PCR amplicifation was in 20 ml using 5 ml template and 0.5 pmol primer for 20 cycles; 1' at 94°C, 1' at 60°C, 1' at 72°C, essentially as described (Craxton
5   1991). Sequence primers: BIGR = 5'AGATTTCCACTTATATCATGTTTTATGTTTTGC (SEQ ID NO:8) , R2 (Van Luenen and Plasterk 1994).

Genetic assay
10   Electrocompetent DS941 lambda lysogen (Flinn et al. 1989) bacteria were prepared and used as described (Zabarovsky and Winberg 1990). The donor plasmid contains a lambda origin of replication and can not replicate in the DS941 lambda lysogen; the target plasmid has Col E1 origin
15   of replication. One to 5 ml of DNA was used per electrophoration and 5% of the bacteria were, after dilution, plated on ampicillin. The remaining bacteria were plated on double selection. This yielded, depending on the efficiency upto 200 transformants.

References

Bainton, R.J., P. Gamas, and N.L. Craig, 1991. Tn7 transposition in vitro proceeds through an excised transposon intermediate generated by staggered breaks in DdNA. *Cell* 65: 805-816.

5   Bainton, R.J., K.M. kubo, J.-n. Feng, and N.L. Craig, 1993.Tn7 transposition: target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in vitro system. *Cell* 72: 931-943.

Beall, E.L., A. Admon, and D.C. Rio, 1994. A Drosophila
10 protein homologous to the human p70 Ku autoimmune antigen interacts with the P transposable element inverted repeats. *Proc.Natl.Acad.Sci.* USA 91: 12681-12685.

Bender, J. and N. Kleckner, 1986. Genetic evidence that Tn10 transposes by a nonreplicative mechanism. *Cell* 45:
15 801-815.

Bolland, S. and N. Kleckner, 1995. The two single-strand cleavages at each end of Tn10 occur in a specific order during transposition. *Proc.Natl.Acad.Sci.* USA 92: 7814-7818.

20 Boyd, A./C. and D.J. Sherratt, 1995. The pCLIP plasmids:versatile cloning vectors based on the bacteriophage lambda origin of replication. *Gene* 153: 57-62.

Clar, J.M. 1988. Novel non-templated nucleotide addition
25 catalyzed by procaryotic and eukaryotic DNA polymerases. *Nucleic Acids Res*. 16: 9677-9686.

Collins, J., B. Saari, and P. Anderson, 1987. Activation of a transposable element in the germ line but not the soma of Caenorhabditis elegans. *Nature* 328: 726-728.

30 Craxton, M, 1991. Linear amplification sequencing, a powerful method for sequencing DNA. METHODS: A companion to Meth. Enzymol. 3: 20-26.

Doak, T.G., F.P. Doerder, C.L. Jahn, and G. Herrick, 1994.
A proposed superfamily of transposase-related genes: new
members in transposon-like elements of ciliated protzoa and
a common "D35E" motif. *Proc.Natl.Acad.Sci.* USA 91: 942-946.

5   Emmons, S.W., L. Yesner, K.-s. Ruan, and D. Katzenberg,
1983. Evidence for a transposon in Caenorhabditis elegans.
Cell 32: 55-65.

Fallaux, F.J., Kranenburg, O., Cramer, S.J., Houweling, A.,
van Ormondt, H., Hoeben, R.C. and van der Eb, A.J. 1996.

10   Characterization of 911: A new helper cell line for the
titration and propagation of early region-1 deleted
adenoviral vectors. Human Gene Therapy 7: 215-222.

Flinn, H.L., M. Burke, C.J. Stirling, and D.J. Sherratt,
1989. Use of gene replacement to construct Escherichia coli

15   strains carrying mutations in 2 genes required for
stability of multicopy plasmids. *J.Bacteriol.* 171: 2241-
2243.

Grindley, N.D.F. and D.J. Sherratt, 1978. Sequence analysis
of IS1 insertion sites: models for transposition. *Cold*

20   *Spring Harbor Symp. Quant. Biol.* 45: 125-133.

Grosschedl, R., K. Giese, and J. Pagel, 1994. HMG domain
proteins: architectural elements in the assembly of
nucleoprotein structures. *Trends Genet.* 10: 94-100.

Henikoff, S, 1992. Detection of Caenorhabditis transposon

25   homologs in diverse organism. *New Biol.* 4: 382-388.

Kaufman, P.D., and D.C. Rio, 1992. P element transposition
in vitro proceeds by a cut-and-paste mechanism and uses GTP
as cofactor. *Cell* 69: 27-39.

Kramer, J.M., R.P. French, E. Park, and J.J. Johnson, 1990.

30   The Caenorhabditis elegans *rol-6* gene, which interacts with
the *sqt-1* collagen gene to determine organismal morphology
encodes a collagen, *Moll.Cel. Biol.* 10: 2081-2089.

Loukeris, T.G., Livadaras, I., Arca, B., Zabalou, S.,
Savakis, C., 1995. Gene transfer into the medfly, Ceratitis

26

capitata, with a Drosophila hydei transposable element. *Science*, 270(5244): 2002-5

Mello, C.C., J.M. Kramer, D. Stinchcomb, and V. Ambros, 1991. Efficient gene transfer in C. elegans: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* 10: 3959-3970.

Mizuuchi, K. 1992. Transpositional recombination: mechanistic insights from studies of Mu and other elements. *Annu. Rev. Biochem.*61: 1011-1051.

Mori, I. 1988. "Analysis of germline transposition and excision of Tcl transposable elements in Caenorhabditis elegans". Ph. D. thesis. Washington University, St. Louis, MO.

Müller, H.-P. and H.E. Varmus, 1994. DNA bending create favored sites for retroviral integration: an explanation for preferred insertion sites in nucleosomes. *EMBO J.* 13: 4704-4714.

Nagai, K. and H.C. Thogersen, 1987. Synthesis and sequence-specific proteolysis of hybrid proteins produced in Escherichia coli. *Methods Enzymol.* 153: 461-481.

Paull, T.T., M.J. Haykinson, and R.C. Johnson, 1993. The nonspecific DNA-binding and -bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structures. *Genes & Dev.* 7: 1521-1534.

Plasterk, R.H.A. 1991. The origin of footprints of the Tcl transposon of Caenorhabditis elegans, *EMBO J.* 10: 1919-1925.

Plasterk, R.H.A., 1995. The Tcl/*mariner* transposon family. In *Transposable Elements* (eds. H. Seadler and A. Gierl), Springer Verlag, Heidelberg.

Pryciak, P.M. and H.E. Varmus, 1992. Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell* 69: 769-780.

Radice, A.D., B. Bugaj, D.H.A. Fitch, and S.W. Emmons, 1994. Widespread occurrence of the Tcl transposon family:

Tc1-like transposons from teleost fish. *Mol. Gen. Genet.*
244: 606-612.

Rio, D.C., G. Barnes, F.A. Laski, J. Rine, and G.M. Rubin,
1988. Evidence for Drosophila P element transposase
activity in mamalian cells and yeast. *J. Mol. Biol.* 200:
411-415.

Robertson, H.M. 1995. The Tc1-mariner superfamily of
transposons in animals. *J.Insect. Physiol.* 41: 99-105.

Robertson, H.M. 1993. The mariner transposable element is
widespread in insects. *Natrure* 362: 241-245.

Robertson, H.M. and D.J. Lmape, 1995. Recent horizontal
transfer of a *mariner* transposable element among and
between Diptera and Neuroptera. *Mol. Biol. Evol.* 12: 850-
862.

Rosenzweig, B., L.W. Liao, and D. Hirsch, 1983. Sequence of
the C. elegans transposable element Tc1. *Nucleic Acids Res.*
12: 4201-4209.

Schukkink, R.F. and Plasterk, R.H.A. 1990. TcA, the
putative transposase of the C.*elegans* Tc1 transposon, has
an N-terminal DNA binding domain. Nucleic Acids Research
18: 895-900.

Shapiro, J.A, 1979. Molecular model for the transposition
and replication of bacteriophage Mu and other transposable
elements. *Proc. Natl. Acad. Sci. USA* 76: 1933-1937.

Van Luenen, H.G.A.M. and R.H.A. Plasterk, 1994. Target site
choice of the related transposable elements Tc1 and Tc3 of
Caenorhabditis elegans. *Nucleic Acids Res.* 22: 262-269.

Van Luenen, H.G.A.M. van Luenen, and R.H.A. Palasterk,
1993. CharActerization of the Caenorhabditis elegans Tc1
transposase in vivo and in vitro. *Genes & Dev.* 7: 1244-
1253.

Vos, J.C. and R.H.A. Plasterk, 1994. Tc1 transposase of
Caenorhabditis elegans is an endonuclease with a bipartite
DNA binding domain. *EMBO J.* 13: 6125-6132.

Way, J.C., L. Wang, J.-O. Run, and A. Wang, 1991. The *mec-3*
gene contains cis-acting elements mediating positive and
negative regulation in cells produced by asymmetric cell
division in Caenorhabditis elegans. *Genes & Dev.* 5: 2199-
5   2211.
Zabarosky, E.R. and G. Winberg, 1990. High efficiency
electrophoration of ligated DNA into bacteria. *Nucleic
Acids Res.* 18: 5912.

Legend

Figure 1. Southern blot analysis of <u>in vitro</u> Tc1

transposition reaction products. Products of <u>in vitro</u>

transposition reactions were separated on a 1% agarose gel,

5   transferred to nitrocellulose and probed with radiolabeled

Tc1. Standard reactions contained Mgcl$_2$ (lanes 1 to 3 and 5

to 11) or EDTA (lane 4). Products were digested with ScaI

in vector DNA (lane 2, 7, 10) or ApaI in Tc1 DNA (lane 3)

prior to electrophoresis. Lanes 5, 8 and 11 show reaction

10   products when the substrate is lenearized with ScaI prior

to in vitro cleavage. Lanes 1 to 5 show reaction products

using pRP466 as substrate which carries a complete Tc1

element (see Figure 4). Lanes 6 to 8 use pRP467 as

substrate which has a deleted left end of Tc1, whereas

15   lanes 9 to 11 show pRP4678 as substrate, which has the

right end of Tc1 deleted. REC and LEC stands for Right and

Left End Cleavage, respectively. RH and LH indicate the

positions of the Right and Left Half of Tc1. A schematic of

ScaI-linearized pRP466 is shown at the bottom of the

20   figure.

Figure 2. Mapping of the <u>in vitro</u> cleavage sites at the

nucleotide level. A PCR based primer extension was

performed on reaction products obtained in the presence of

Mgcl$_2$ (+) or EDTRA(-) using pRP466 as donor. A control

25   reaction was performed with pRP466 digested with EcoRV (RV

lanes) to demonstrate the addition of one extra nucleotide

at the end of the PCR product by Taq polymerase (see Clark

1988). Products were analyzed on a sequencing gel. Sequence

reactions (GATC) were loaded as markers. PCR was with

30   primer R2 (right panel) or primer BIGR (left panel). The

relevant sequence is indicated with the EcoRV site boxed

and the TA target site underlined. The cleavage sites are

shown by arrows. Identical results were obtained when the

positions of cleavage at the other transposon end was

35   determined.

Figure 3. Schematic representation of the genetic transposition assay. The donor plasmid, a pACB104 derivative (Boyd and Sherratt 1995) with a lambda origin of replication, contains a Tc1 element carrying an antibiotic

5 resistance gene. The target plasmid, pRP475, carries a 1,4 kb *hind*II *gpa-2* fragment and a Col E1 origin of replication (pSP72, Promega). Reaction products were electroporated into a lambda lysogen. *E. coli* strain to counterselect against the donor. Integration events were selected on

10 double antibiotics.

Figure 4. Target site selection. Comparison of the distribution of <u>in vitro</u> (black bars) and <u>in vivo</u> (open bars) Tc1 insertions. pRP472 as donor and pRP475 as target were used in standard <u>in vitro</u> transposition reactions

15 using *C. elegans* extract. Every mark on the X-axis represents a TA dinucleotide in the *gpa-2* fragment as described in detail elsewhere (Van Luenen and Plasterk 1994).

Figure 5. Purification of Tc1 transposase from *E. coli*.

20 Analysis of transposase purified from inclusion bodies on a 12% SDS-polyacrylamide gel. Lane M: molecular weight markers (indicated in KDa); lane 1: bacterial lysate before induction; lane 2: bacterial lysate after induction: lane 3: purified inclusion bodies; lane 4: purified transposase

25 after refolding.

Figure 6. Mutations at the extreme termini of Tc1 affect excision. In vitro reaction products w3ere obtained using *C. elegans* extract in the presence of Mgcl$_2$ (+) or EDTE (-) using as donor pRP480 (wt), pRP481 (TA), pRP482 (CA),

30 pRP483 (GT) or pRP484 (BS), as indicated at the top. Products were separated on a 1% agarose gel, transferred to nitrocellulose and probed with radiolabeled Kan[R]-gene fragment.The donor plasmids contain 28-mers cloned into the *Sma*I-site (wt sequence) and the *Hind*II-site (wt or

35 mutant sequence) of pUC19 with the Kan[R]-cassette of pUC4K

in between. TA was mutated to CG, CA to TG and GT to AC,
respectively. In the transposase binding site mutation the
*Bal*I and *Eco*RV sites are mutated to TCCCA and GGGCCC,
respectively (see Vos and Plasterk 1994).

5   **Figure 7.** Model for Tc*1* transposition. A model for non-
replicative Tc*1* transposition showing the excised, linear
element with a 2 bp 3' staggered overhang. Integration
results in a duplication of the TA target site. Repair of
the double strand break leads to the generation of

10   characteristic footprints (see also Van Luenen et al.
1994).

Figure 8: Tc1-transposase is not toxic for human cells.
The cell line 911 was transfected with 2 µg pRc/CMV.Tc1A or
2 µg pRc/CMV per 5 cm, and monoclonal cell lines were

15   established from G418 resistant colonies. From each cell
line, 50 µg whole cell extract was separated on a 10% PAA
gel and subsequently analysed for Tc1-transposase. Tc1-
transposase was detected in the cell lines transfected with
pRc/CMV.Tc1A (lanes 1-5), whereas no Tc1-transposase could

20   be detected in the negative control cell lines transfected
with pRc/CMV (lanes 6-8).

Figure 9: Southern blot of *in vitro* excision reaction
products of pRP466.SV-neo1.PGK-tk2, catalysed by Tc1-
transposase purified from E.*coli*.. Incubation of the donor

25   plasmid pRP466.SV-neo1.PGK-tk2 with purified Tc1-
transposase (lane 1), but not with NEB (lane 2) results in
excision of the Tc1.SV-neo transposon (lane 1). Thus, the
plasmid pRP466.SV-neo1.PGK-tk2 can serve as a Tc1.SV-neo
transposon donor plasmid.

30

Table 1. Transposition frequencies. In vitro Tc*1*
transposition reactions were carried out with supercoiled
(SC) or linear donor plasmids and with protein sources as
indicated and the ratios of amp^R-kan^R to amp^R colonies

35   (*10^6) are shown for two independent experiments. No

integration products were recovered when reactions were
performed in the presence of EDTA.

Table 2: Formation of G418 resistant colonies after co-
transfection of pRP466.SV-neol.PGK-tk2 with pcDNA1/Tc1A or
5　pcDNA1/amp. The cell line 911 was co-transfected with
pRP466.SV-neol.PGK-tk2 and pcDNA1/Tc1A or pcDNA1/amp. For a
period of 18 days, the transfected cells were put on G418
selection medium . Subsequently, the G418 resistant
colonies were stained with methylene blue and counted. Co-
10　transfection of pRP466.SV-neol.PGK-tk2 with pcDNA1/Tc1A
yielded approximately 40% more G418 resistant colonies than
co-transfection with pcDNA1/amp. This indicates that , in
addition to random integration, G418 resistant colonies
were formed due to jumping of the Tc1-neo transposon into
15　the cellular genome.

Table 3: Expression of HSV-tk in G418 resistant cell lines
co-transfected with pRP466.sv-neol.PGK-tk2 with pcDNA1/Tc1A
or pcDNA1/amp. Co-transfection of pRP466.sv-neol.PGK-tk2
with pcDNA1/Tc1A yielded more HSV-tk negative (Ganciclovir
20　resistant) colonies than co-transfection with pcDNA1/amp.

Table 1

Donor	source	exp.1	exp.2
pRP490, SC	C. elegans	21	22
pRP490, linear	C. elegans	0.5	1.0
pRP491, SC	C. elegans	3.7	1.6
pRP490, SC	E. coli	3.0	3.2

Table 2

Transfected DNA/5 cm dish	n=	mean number of colonies
2 μg pcDNA1/Tc1A + 0.1 μg pRP466.SV-neo1.PGK-tk2	3	51
2 μg pcDNA1/amp + 0.1 μg pRP466.SV-neo1.PGK-tk2	3	37
2 μg pcDNA1/Tc1A + 0.5 μg pRP466.SV-neo1.PGK-tk2	2	206
2 μg pcDNA1/amp + 0.5 μg pRP466.SV-neo1.PGK-tk2	2	144
2 μg pcDNA1/Tc1A + 2 μg pRP466.SV-neo1.PGK-tk2	2	342
2 μg pcDNA1/amp + 2 μg pRP466.SV-neo1.PGK-tk2	2	251
no DNA	2	0

Table 3

Transfected DNA	n=	% HSV-tk negative
2 µg pcDNA1/Tc1A + 0.1 µg pRP466.SV-neo1.PGK-tk2	16	77
2 µg pcDNA1/Tc1A + 0.5 µg pRP466.SV-neo1.PGK-tk2	19	64
2 µg pcDNA1/amp + 0.1 µg pRP466.SV-neo1.PGK-tk2	19	32
2 µg pcDNA1/amp + 0.5 µg pRP466.SV-neo1.PGK-tk2	18	44

SEQUENCE LISTING

(1) GENERAL INFORMATION:

   (i) APPLICANT:
     (A) NAME: Valerio, Domenico
     (B) STREET: Gerbrandylaan 12
     (C) CITY: Leiden
     (D) STATE: Zuid-Holland
     (E) COUNTRY: the Netherlands
     (F) POSTAL CODE (ZIP): 2314 EZ

     (A) NAME: Schouten, Govert Jan
     (B) STREET: Da Costastraat 82
     (C) CITY: Leiden
     (D) STATE: Zuid-Holland
     (E) COUNTRY: the Netherlands
     (F) POSTAL CODE (ZIP): 2321 AR

     (A) NAME: Plasterk, Ronald Hans Anton
     (B) STREET: Zwarteweg 5
     (C) CITY: Bussum
     (D) STATE: Noord-Holland
     (E) COUNTRY: the Netherlands
     (F) POSTAL CODE (ZIP): 1405 AA

     (A) NAME: van Luenen, Hendricus Gerhard Adrianus Maria
     (B) STREET: Leerdamhof 143
     (C) CITY: Amsterdam
     (D) STATE: Noord-Holland
     (E) COUNTRY: the Netherlands
     (F) POSTAL CODE (ZIP): 1108 BM

  (ii) TITLE OF INVENTION: Vectors and methods for providing cells with
     additional nucleic acid material integrated in the genome
     of said cells.

 (iii) NUMBER OF SEQUENCES: 12

  (iv) COMPUTER READABLE FORM:
     (A) MEDIUM TYPE: Floppy disk
     (B) COMPUTER: IBM PC compatible
     (C) OPERATING SYSTEM: PC-DOS/MS-DOS
     (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

  (v) CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US 08/909,786

(2) INFORMATION FOR SEQ ID NO: 1:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 44 base pairs
     (B) TYPE: nucleic acid
     (C) STRANDEDNESS: unknown
     (D) TOPOLOGY: unknown

  (ii) MOLECULE TYPE: other nucleic acid

 (iii) HYPOTHETICAL: NO

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CCCCAAGCTT GCCACCATGG TAAAATCTGT TGGGTGTAAA AATC
44

(2) INFORMATION FOR SEQ ID NO: 2:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 32 base pairs
     (B) TYPE: nucleic acid
     (C) STRANDEDNESS: unknown
     (D) TOPOLOGY: unknown

  (ii) MOLECULE TYPE: other nucleic acid

 (iii) HYPOTHETICAL: NO

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GCTCTAGATG CTTAATACTT TGTCGCGTAT CC   32

(2) INFORMATION FOR SEQ ID NO: 3:

   (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 56 base pairs
     (B) TYPE: nucleic acid
     (C) STRANDEDNESS: unknown
     (D) TOPOLOGY: unknown

  (ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GATCCAAGGA GAGGACGCTG TCTGTCGAAG GTAAGGAACG
GACGAGAGAA GGGAGA   56

(2) INFORMATION FOR SEQ ID NO: 4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 52 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: other nucleic acid

    (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TCTCCCTTCT CGAATCGTAA CCGTTCGTAC GAGAATCGCT GTCCTCTCCT
TG   52

(2) INFORMATION FOR SEQ ID NO: 5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: other nucleic acid

    (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TCAAGTCAAA TGGATGCTTG AG   22

(2) INFORMATION FOR SEQ ID NO: 6:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 24 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: unknown
      (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: other nucleic acid

  (iii) HYPOTHETICAL: NO

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GATTTTGTGA ACACTGTGGT GAAG   24

(2) INFORMATION FOR SEQ ID NO: 7:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 20 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: unknown
      (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: other nucleic acid

  (iii) HYPOTHETICAL: NO

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GTACGAGAAT CGCTGTCCTC   20

(2) INFORMATION FOR SEQ ID NO: 8:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 33 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: unknown
      (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: other nucleic acid

  (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

AGATTTCCAC TTATATCATG TTTTATGTTT TGC   33

(2) INFORMATION FOR SEQ ID NO: 9:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: other nucleic acid

    (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TCCCA   5

(2) INFORMATION FOR SEQ ID NO: 10:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 6 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: other nucleic acid

    (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GGGCCC   6

(2) INFORMATION FOR SEQ ID NO: 11:

    (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CTCTGCTCAT ATGTCACGAC CGGTTTTTCT ATAGGTG   37

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CACCTATAGA AAAACCGGTC GTGACATATG AGCAGAG
37

CLAIMS

1.   A vector for integrating additional nucleic acid
material into a cell of a certain genome of a first genus,
said vector comprising:
     two transposase binding sites and
5   a cut site corresponding to each of said transposase
binding sites, wherein said transposase binding sites are
derived from a transposon found in a second genus;
     wherein each of said transposase binding sites are in
close proximity to said corresponding cut site; and
10   wherein the additional nucleic acid material is
integrated between said two transposase binding sites of
the vector.

2.   A vector according to claim 1, whereby the transposase
binding sites and cut sites are based on the corresponding
15 sites in transposons from the Tc1/mariner superfamily of
transposons.

3.   A vector according to claim 1 or 2, whereby the
transposase binding sites and cut sites are based on Tc1-
like transposons.

20 4.   A vector according to claim 3, comprising at least the
terminal 26 basepairs of Tc1 as a transposase binding site
and a cut site for the Tc1 transposase.

5.   A vector according to anyone of the aforegoing claims,
further comprising a means for transducing a target cell.

25 6.   A vector according to anyone of the aforegoing claims
further comprising a nucleic acid sequence encoding
transposase activity, functional for the transposase
binding sites and cut sites present in said vector.

7.   A vector according to anyone of the aforegoing claims,
30 whereby the additional nucleic acid material encodes a
protein.

8.   A vector according to anyone of claims 1-7, whereby
the additional nucleic acid material provides a blocking
nucleic acid, such as an antisense RNA.

9.   A vector according to anyone of the aforegoing claims,
whereby said vector is derived from a viral vector,
preferably an adenoviral, retroviral, or adeno associated
viral vector.

5   10.   A method for providing a target cell with additional
nucleic acid material integrated into its genome,
comprising transducing said cell with a vector according to
anyone of the aforegoing claims and providing said target
cell with transposase activity functional for the

10   transposase binding sites and cut sites present on said
vector.

11.   A method according to claim 10, whereby the
transduction is carried out through infectious particles
comprising the vector.

15   12.   A recombined target cell obtainable by a method
according to claim 10 or 11.

13.   Use of a functional transposon found in a certain
genus in integrating a nucleic acid of interest into the
genome of a cell of a different genus.

20   14.   Use according to claim 13, in a gene therapy regime.

15.   Use according to claim 13 for producing transgenic
animals.

Fig. 1
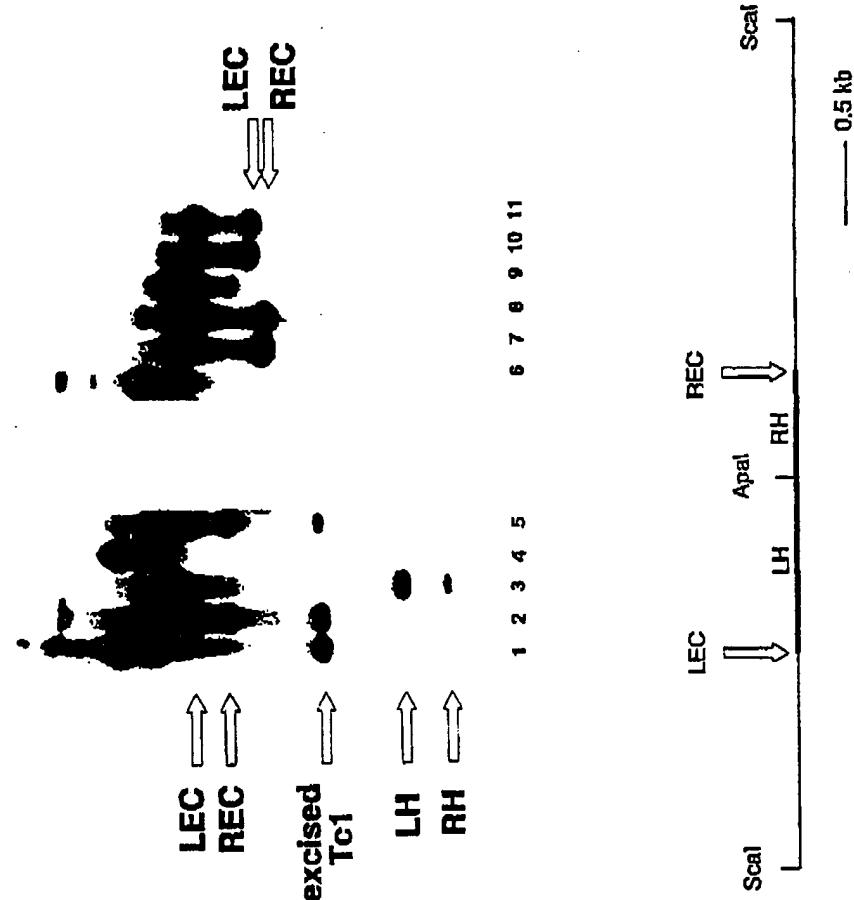
Fig. 2
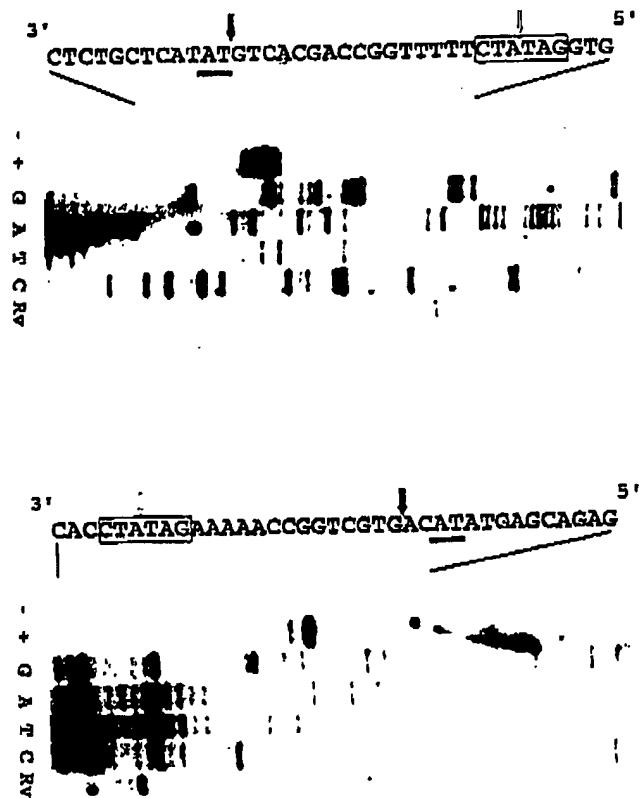
Fig. 3
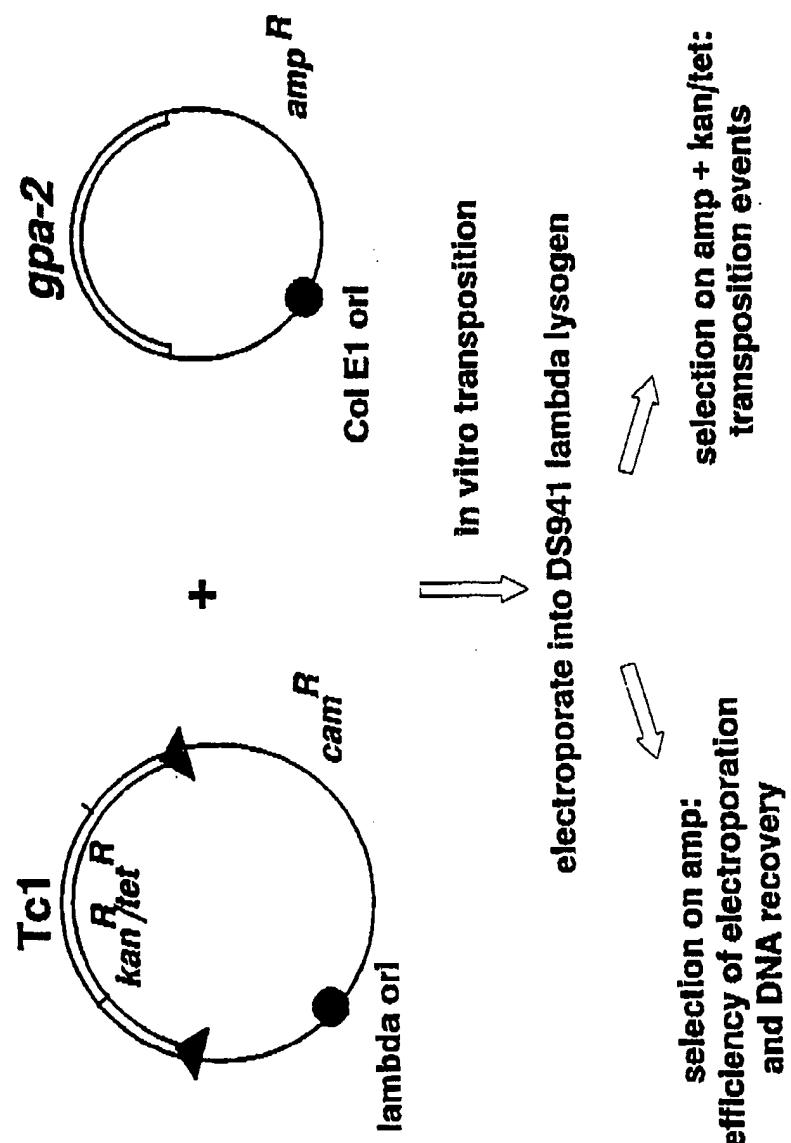
Fig. 4
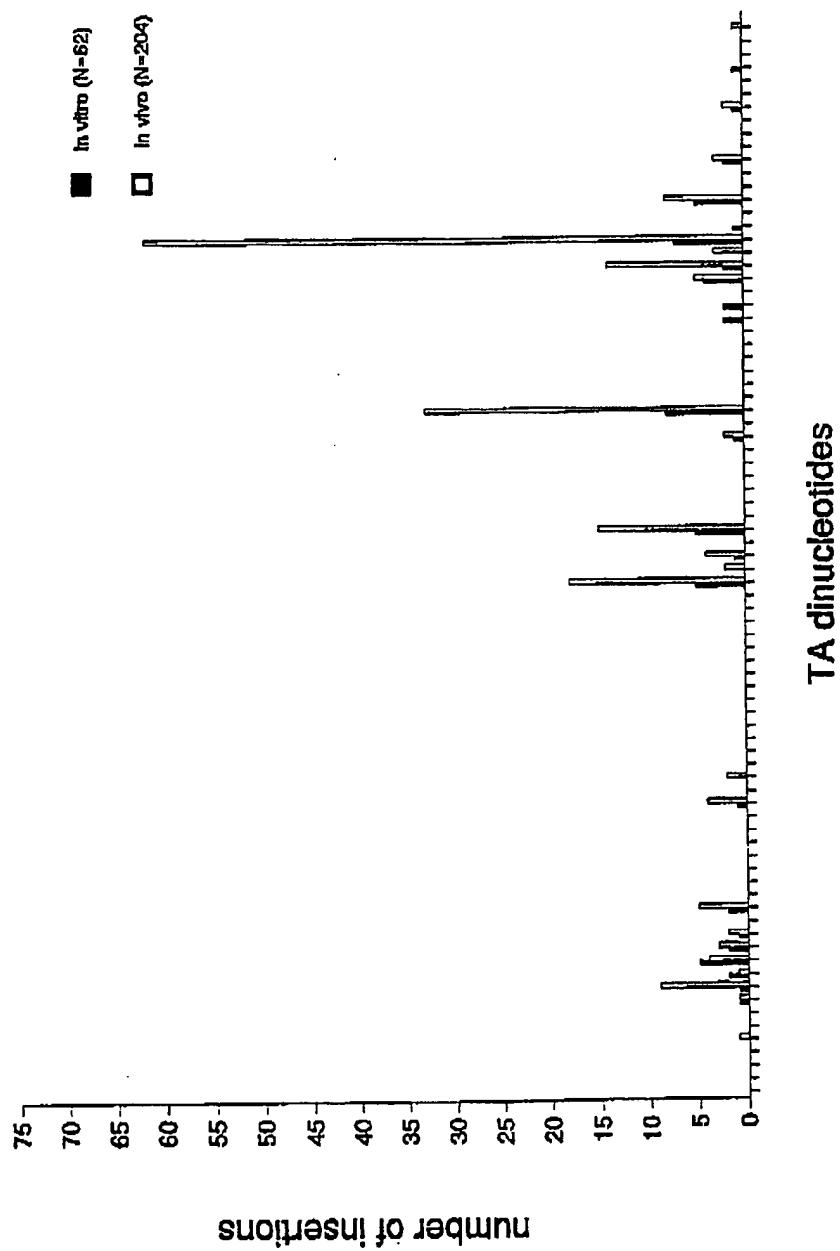
Fig. 6
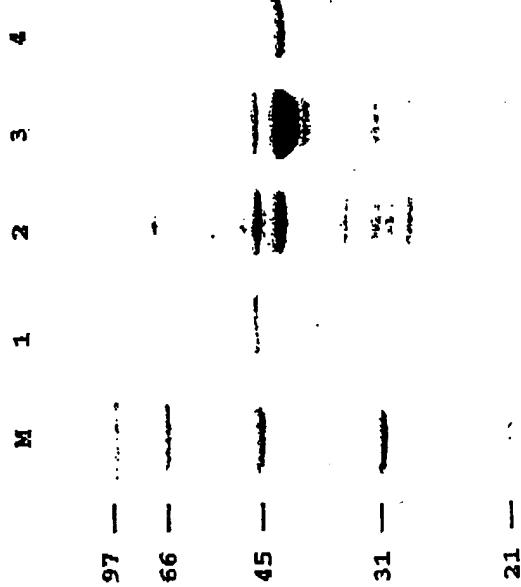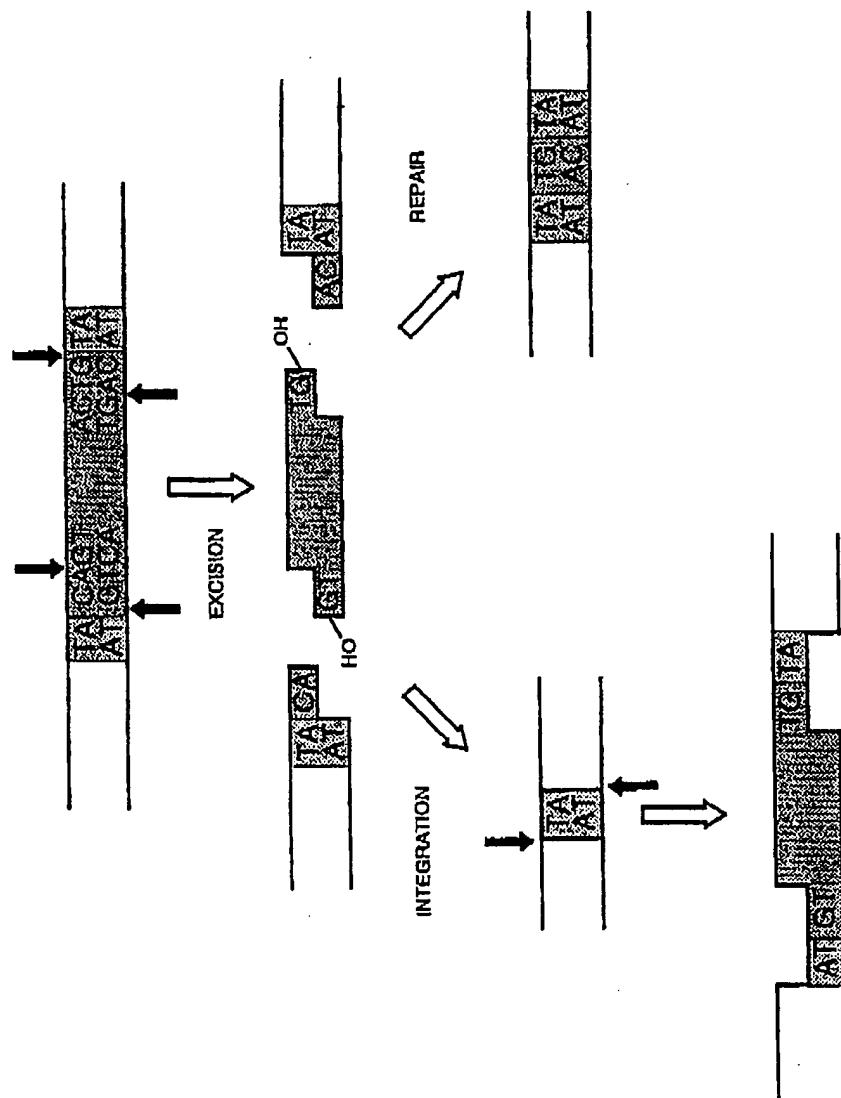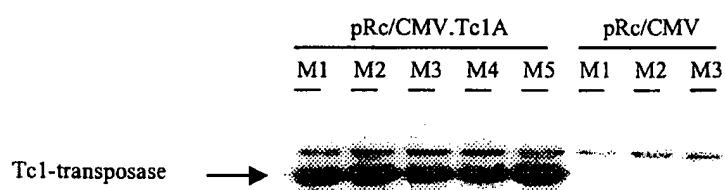
Fig. 7

Figure 8:

Figure 9:

$\underline{+}$ $\underline{-}$ purified transposase

excised Tc1.SV-neo   ⟶
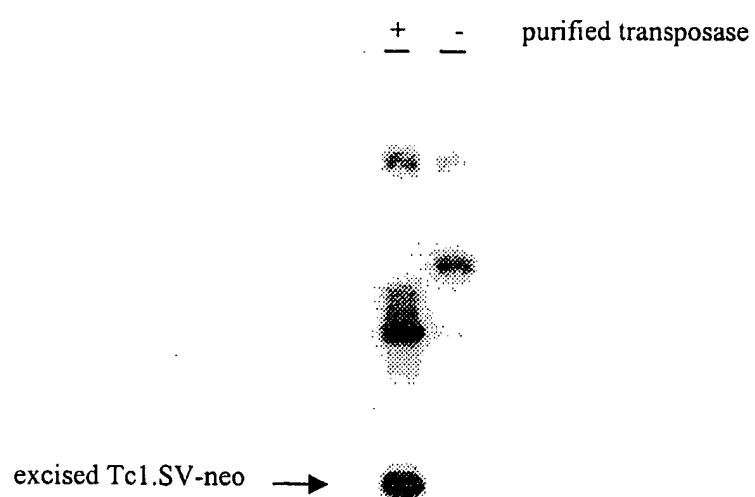
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6   C12N15/90   C12N15/55   C12N9/22   C07K14/435   A01K67/027
        A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6   C12N   C07K   A01K   A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,L, X	WO 97 29202 A (HET NL KANKER INSTITUUT ;PLASTERK RONALD HANS ANTON (NL); VOS JAN) 14 August 1997 see the whole document ---	1-15
X	J.C. VOS ET AL.: "Transposase is the only nematode protein required for in vitro transposition of TC1" GENES & DEVELOPMENT, vol. 10, no. 6, 15 March 1996, pages 755-761, XP002088205 CSH LABORATORY PRESS, NEW YORK, US see the whole document --- -/--	1-8,10, 12,13

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 December 1998	25/01/1999

Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Hornig, H

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X	SCHOUTEN G J ET AL: "Transposon Tc1 of the nematode Caenorhabditis elegans jumps in human cells." NUCLEIC ACIDS RESEARCH, (1998 JUN 15) 26 (12) 3013-7. JOURNAL CODE: O8L. ISSN: 0305-1048., XP002088087 see the whole document	1-15
A	VOS J C ET AL: "CHARACTERIZATION OF THE CAENORHABDITIS ELEGANS TC1 TRANSPOSASE IN VIVO AND IN VITRO" GENES AND DEVELOPMENT, vol. 7, no. 7A, 1 July 1993, pages 1244-1253, XP000575689 see the whole document	1-15
A	VOS J C ET AL: "TC1 TRANSPOSASE OF CAENORHABDITIS ELEGANS IS AN ENDONUCLEASE WITH A BIPARTITE DNA-BINDING DOMAIN" EMBO JOURNAL, (15 DEC 1994) VOL. 13, NO. 24, PP. 6125-6132. ISSN: 0261-4189., XP002088088 cited in the application see the whole document	1-15
A	H.G.A.M. VAN LUENEN AND R.H.A. PLASERK: "Target site choice of the related transposable elements Tc1 and Tc3 of Caenorhabditis elegans" NUCLEIC ACIDS RESEARCH, vol. 22, no. 3, 1994, pages 262-269, XP002088089 IRL PRESS LIMITED,OXFORD,ENGLAND cited in the application see the whole document	1-15
A	ABAD P ET AL: "Expression and biochemical characterization of the DNA binding activity of TcA, the putative transposase of Caenorhabditis elegans transposable element Tc1." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, (1993 MAY 14) 192 (3) 1445-52. JOURNAL CODE: 9Y8. ISSN: 0006-291X., XP002088090 see the whole document	1-15

2

INTERNATIONAL SEARCH REPORT

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9729202 A	14-08-1997	AU 1674697 A	28-08-1997